

Autonomous Vocal and Backing Track Mixing

Master Project Proposal

Kelian (Mike) Li

Music Informatics Group

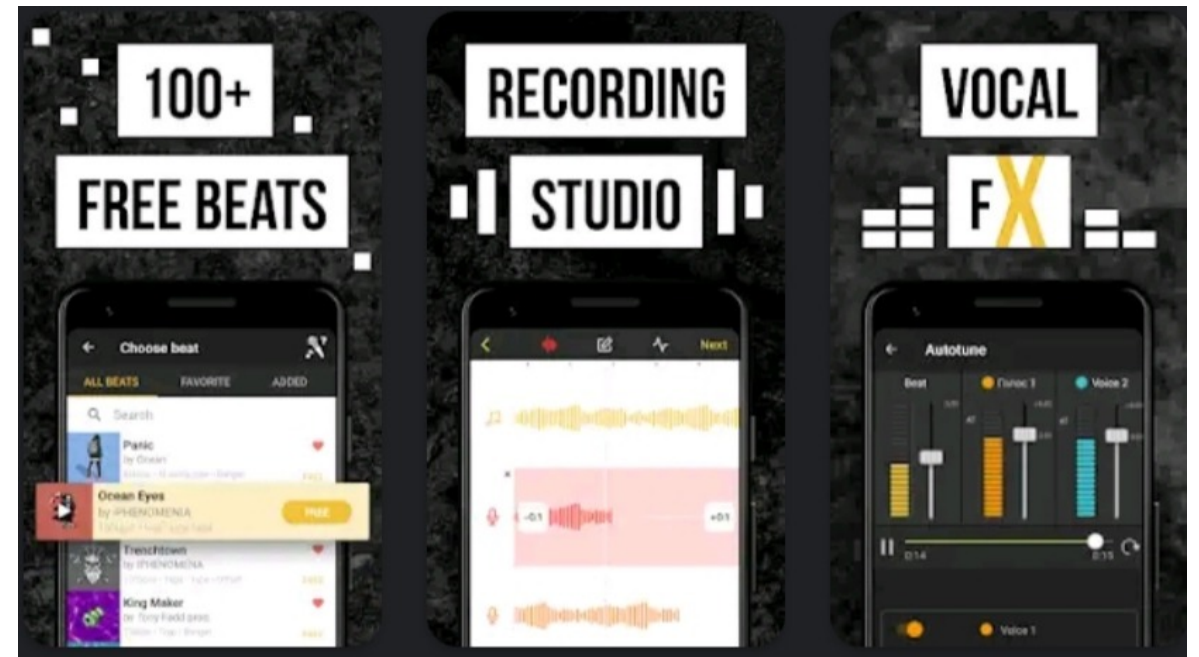
Motivation

- Karaoke apps



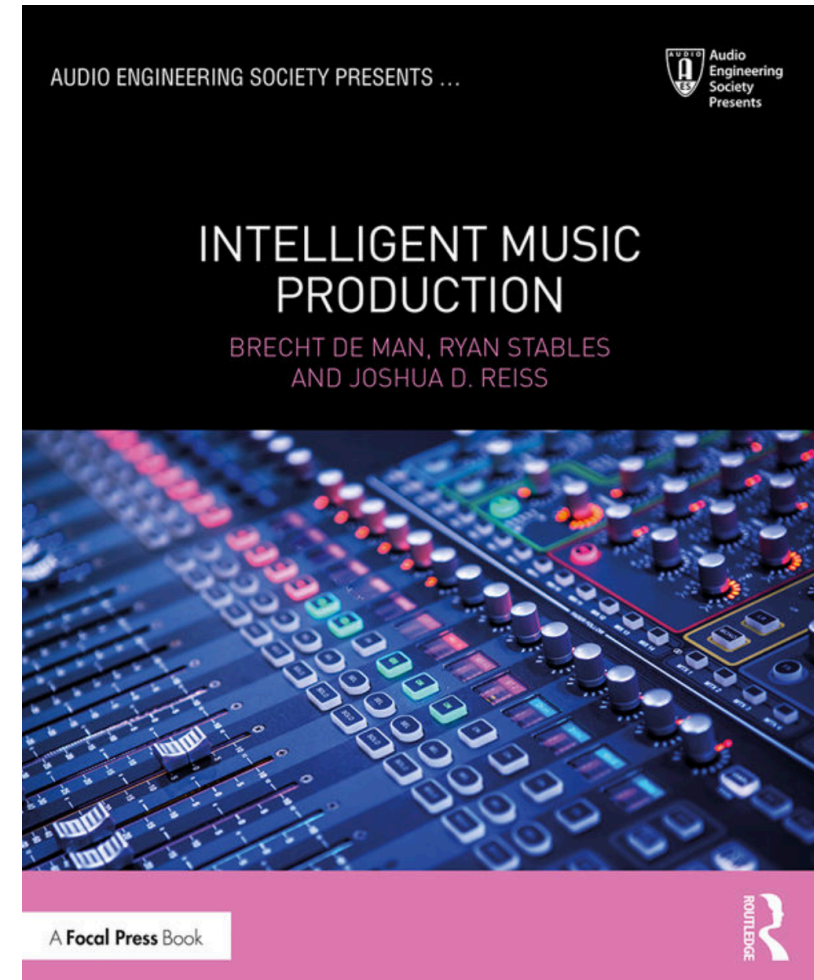
Motivation

- Karaoke apps
- Amateur music makers



Related Works

- Over 10 years of study on automatic mixing
- Mixing rules and mix analysis



Baseline system

- Level balance
 - -3 dB vocal-to-mix ratio
- Compression
 - 14 dB loudness range
- EQ
 - Frequency masking
- Reverb
 - ?

Related Works

- Recent deep learning methods
 - End-to-end audio transformation
 - Differentiable digital signal processing
 - Black-box audio effects control with gradient approximation
- Lack of **paired raw tracks and mixed tracks** for training

Proposed Method

- Maps the input audio to mixing parameters
 - Outputs mixing parameters which allows human adjustment
 - The challenge is the mixing parameter gathering
- Requires only **mixed vocal and backing tracks** for training
 - Raw vocal tracks are not needed

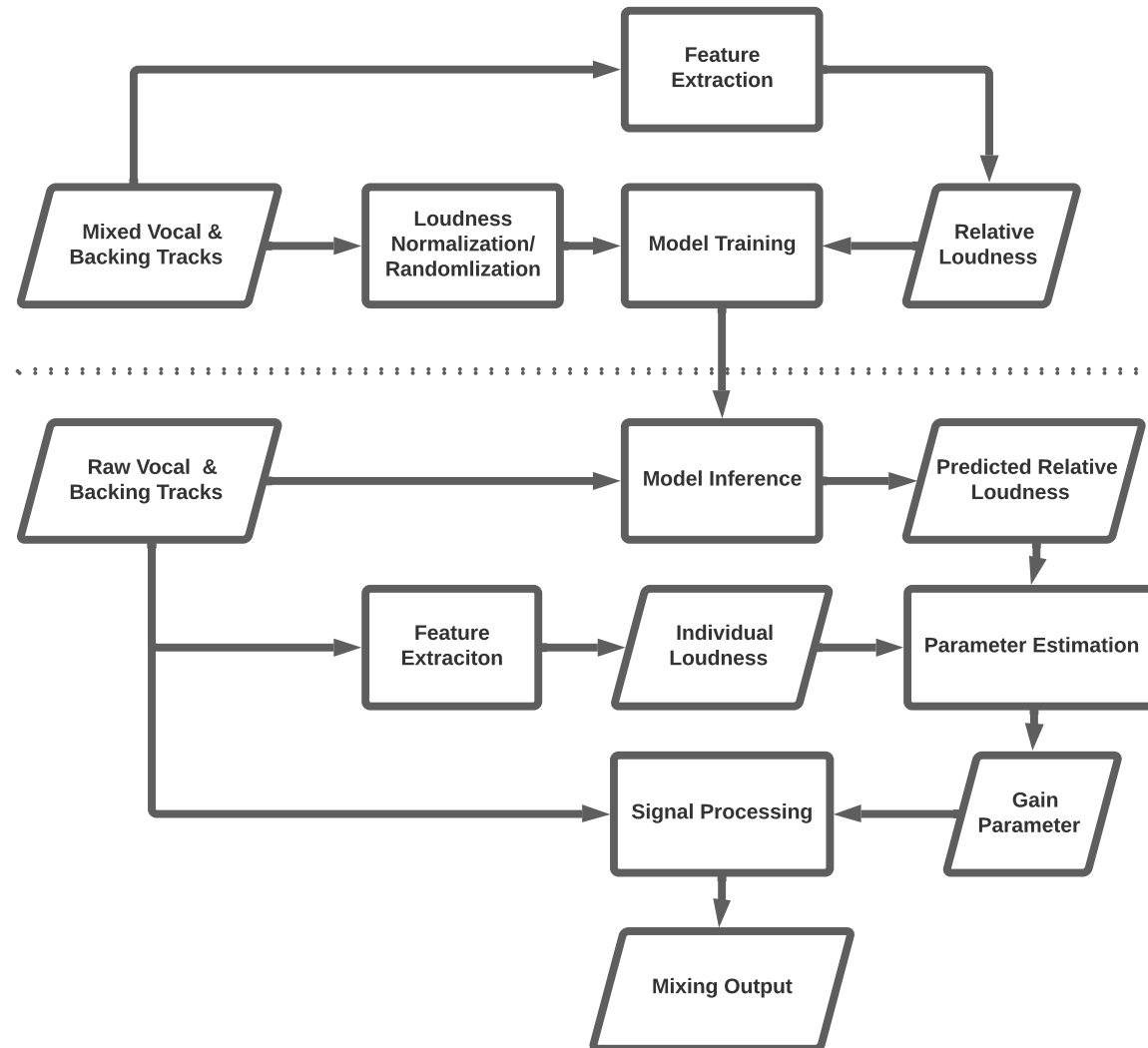
Proposed Method

-Level Balance and Compression

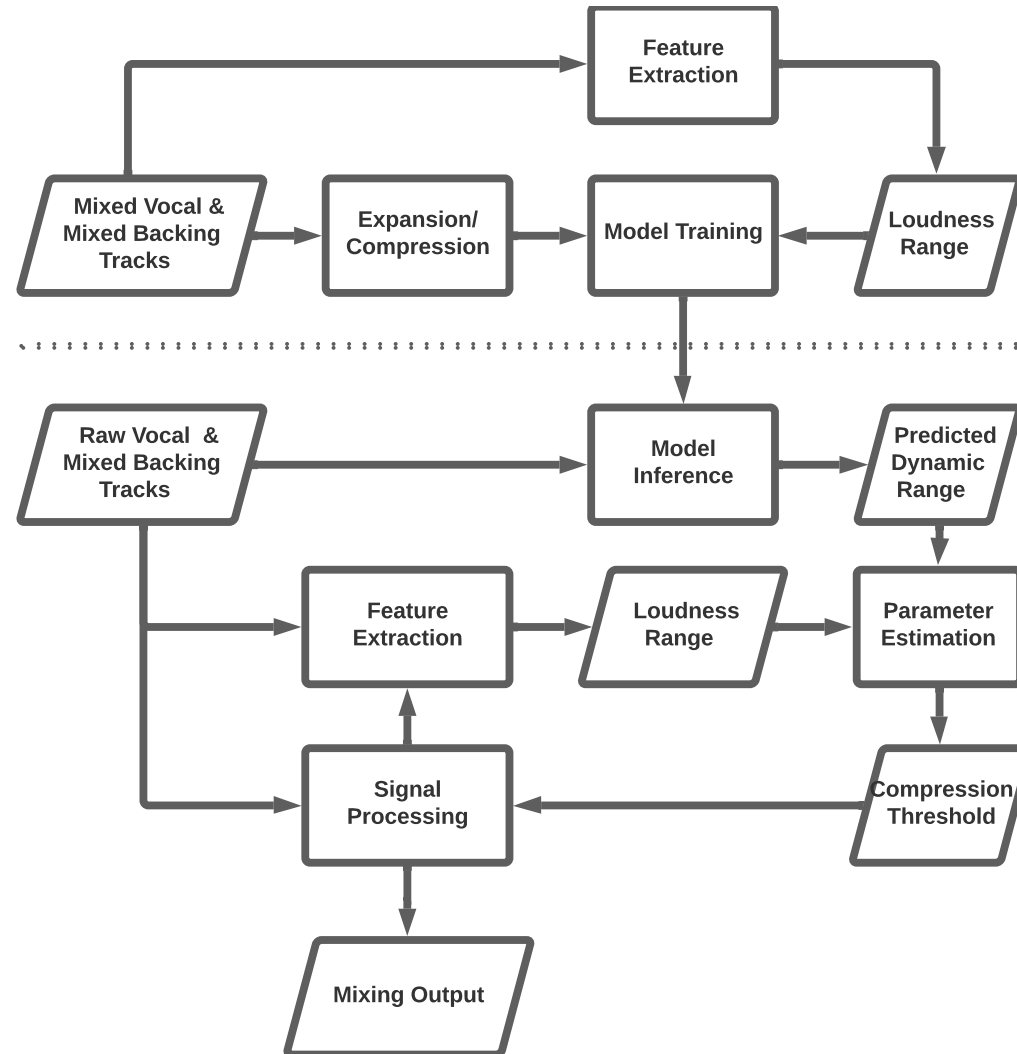
- The model outputs intermediate audio features (relative loudness and loudness range)
- Post-processing converts the intermediate features into mixing parameters

Proposed Method

-Level Balance



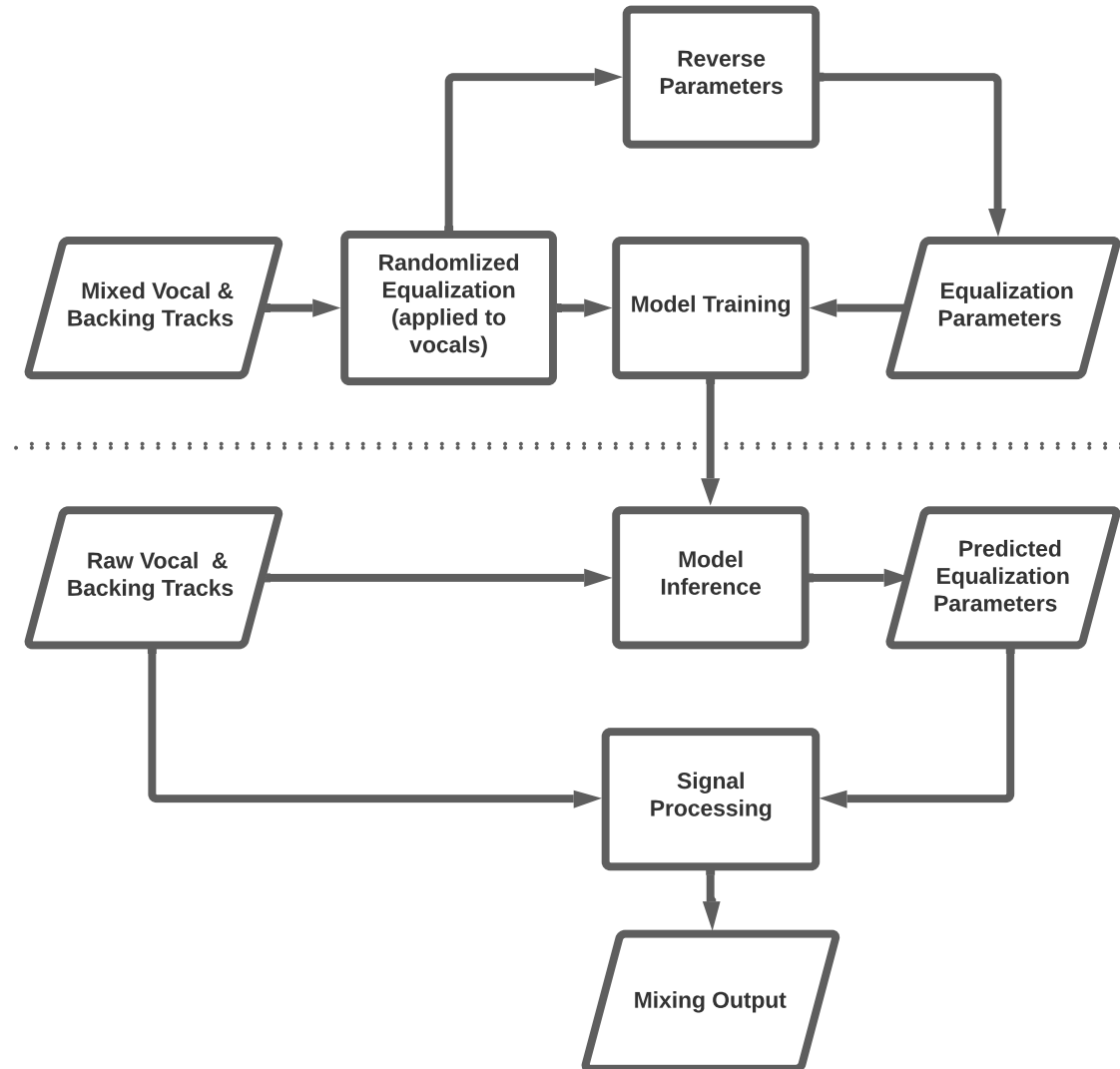
Proposed Method -Compression



Proposed Method -Equalization

- The “raw” tracks for training are **self-generated** by applying EQ to the mixed vocal tracks. The corrected parameters are known.

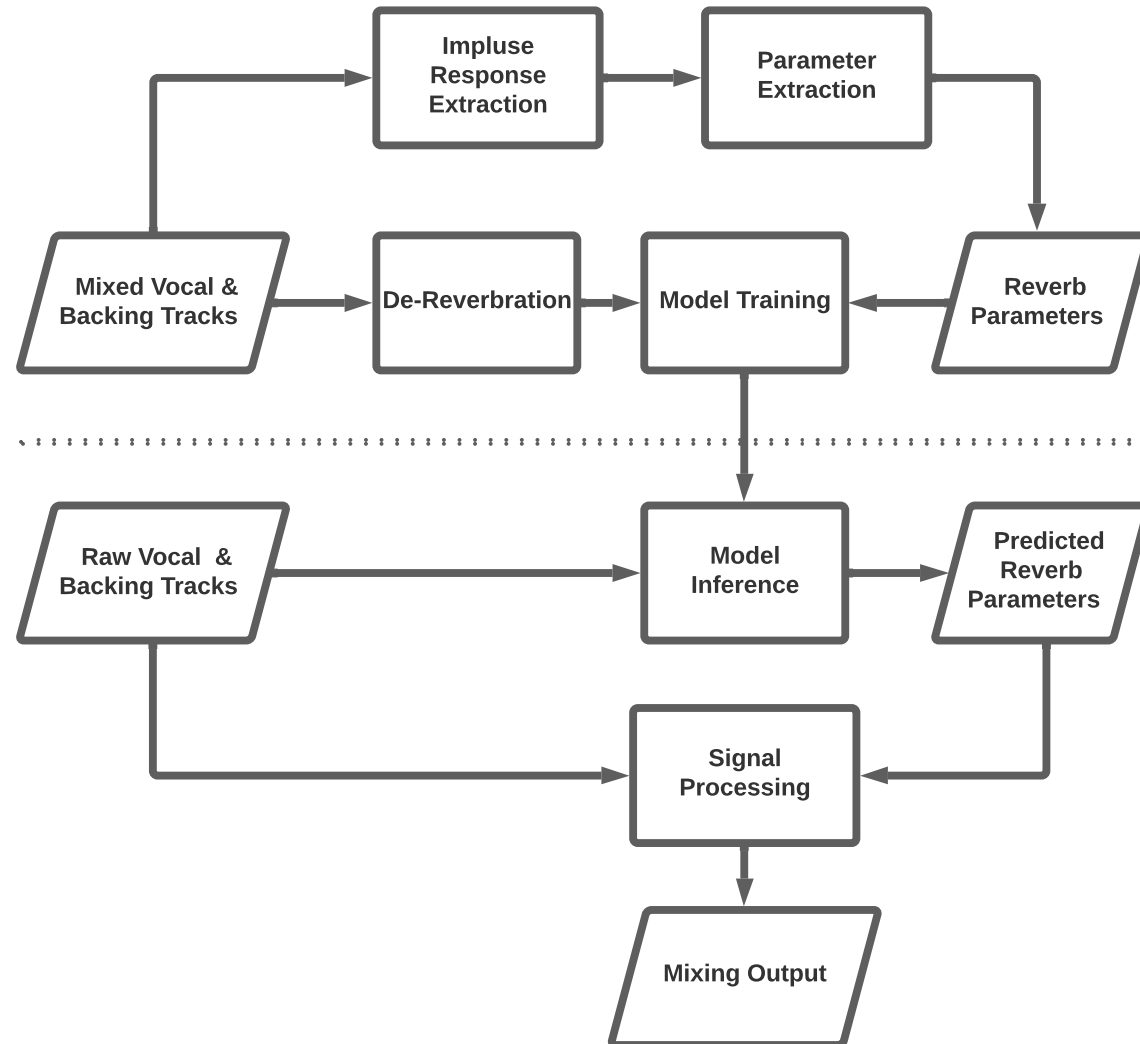
Proposed Method -Equalization



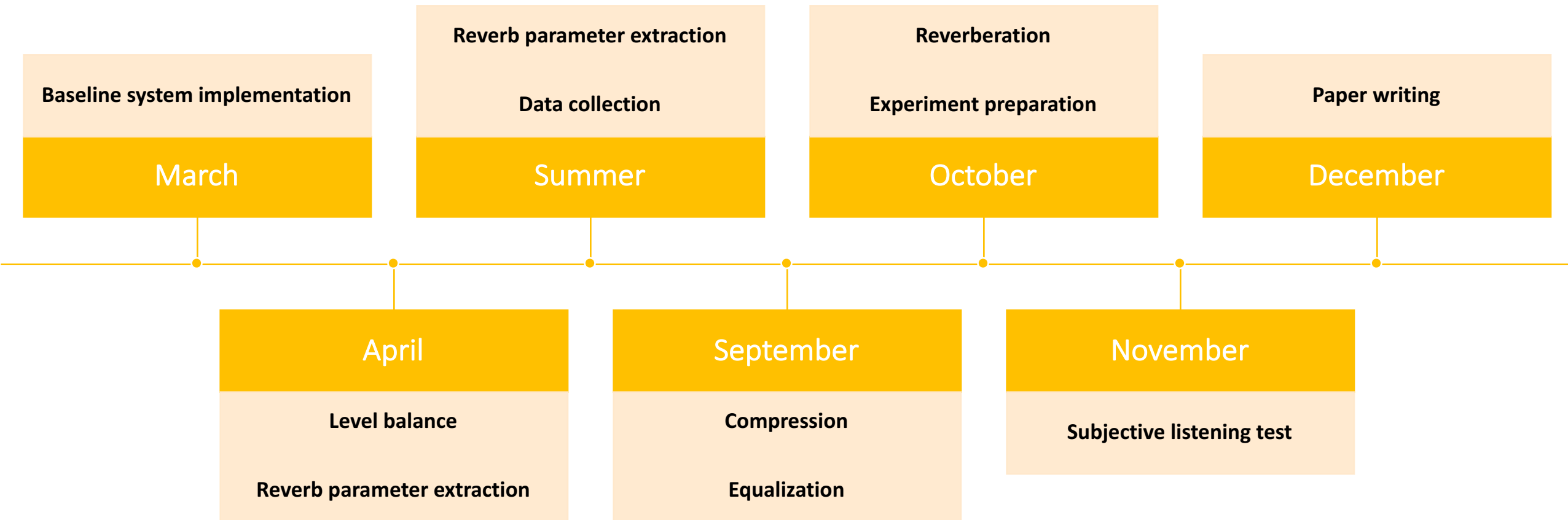
Proposed Method -Reverberation

- Extracts the reverb impulse responses by a commercial plugin
- Uses genetic optimization to approximate the reverb parameters for the impulse responses

Proposed Method -Reverberation



Timeline



Thank you!